

# Robust GMM analysis of models for the short rate process

Rosario Dell'Aquila<sup>a,b,\*</sup>, Elvezio Ronchetti<sup>b,c</sup>, Fabio Trojani<sup>b</sup>

<sup>a</sup> *Quantitative Investment Research, Zürcher Kantonalbank, Zurich, Switzerland*

<sup>b</sup> *Institute of Finance, University of Southern Switzerland, Via Buffi 13, CH-6900 Lugano, Switzerland*

<sup>c</sup> *Department of Econometrics, University of Geneva, Blv. Pont d'Arve 40, CH-1211 Geneva, Switzerland*

---

## Abstract

We re-examine the empirical evidence concerning a well-known class of one-factor models for the short rate process (cf. Chan et al. [Journal of Finance 47 (1992) 1209] (CKLS)) and some recent extensions allowing for a nonlinear drift and for changing parameters with a new statistical methodology based on robust statistics, the Robust Generalized Method of Moments (RGMM). We find that standard GMM model selection procedures are highly unstable in these applications. When testing the CKLS models with the RGMM we find that they are all clearly misspecified and we identify a clustering of influential observations in the 1979–1982 subperiod, a time span that is well known to coincide with a temporary change in the monetary policy of the Federal Reserve. This clustering of influential observations does not disappear when we introduce a non-linearity in the drift and allow for a parameter shift during the 1979–1982 period. Moreover, a Cox–Ingersoll–Ross model (selected by the RGMM) might offer a satisfactory data description for the period after 1982, since there only a few isolated outliers are found. Comparable results are obtained for the Euro-mark case.

© 2002 Elsevier Science B.V. All rights reserved.

*JEL classification:* C1; C5; E4

*Keywords:* GMM estimators and tests; One-factor models of interest rates; Robust estimation; Robust testing; Robust model selection

---

## 1. Introduction

This paper deals with the application of robust statistics in finance and more precisely with the problem of robustly estimating and testing models for the short rate process.

---

\* Corresponding author. Quantitative Investment Research, Zürcher Kantonalbank, Zurich, Switzerland.

*E-mail addresses:* rosario.dellaquila@lu.unisi.ch (R. Dell'Aquila), elvezio.ronchetti@metri.unige.ch (E. Ronchetti), fabio.trojani@lu.unisi.ch (F. Trojani).

The theory of robust statistics is concerned with the construction of statistical procedures that are stable even when the underlying model is not perfectly satisfied by the available dataset (see [Huber, 1981](#); [Hampel et al., 1986](#) for an overview). It can deal with a part of the data that is not fully compatible with the distributions implied by the assumed model and therefore can be seen as a statistical theory dealing with approximate models, rather than with perfectly specified ones.

Many classical econometric procedures are well-known for not being robust, because their results may depend crucially on the properties of a few observations in the sample and in the extreme case on the properties of only one of them. These procedures are optimal when the assumed model is exactly satisfied, but they are biased and/or inefficient when some (even small) model deviations are present. The results obtained by classical procedures can therefore be misleading in such situations.

Financial models are often estimated and tested with methodologies that do not explicitly control for the effects of small distributional deviations from the assumptions.<sup>1</sup> However, because of the intrinsic complexity of financial markets and the richness of financial phenomena, we may realistically believe that some deviations from the assumptions will (almost) always be present when using a financial model in empirical finance.

It seems therefore natural to treat financial models as approximate descriptions of the financial reality and to work with statistical procedures that can deal with some amount of “abnormal” observations and identify them systematically. In some cases, it is precisely a detailed analysis of the identified abnormal observations that will offer new insights and suggestions on the kind of features that a more accurate model should be able to fit.

Implicitly, we argue that while estimating a financial model it is important to verify first if the *majority* of the data is consistent with the assumed model. If this is not the case, a more complex model can be introduced. This seems particularly meaningful in the context of empirical financial modelling, where parameter estimates and the model selected are often the input for the pricing and hedging of financial instruments. In practice, one would like to ensure that the choice of a model used to price and hedge a financial instrument is driven by the features of the majority of the observed data rather than by single datapoints or some particular historical period.

Our major purpose in this paper is methodological. We illustrate the robustness problem by using a robust version of the Generalized Method of Moment (GMM) (hereafter RGMM, see [Ronchetti and Trojani, 2001](#)) within well-known financial applications of the GMM and demonstrate how the RGMM may be used to produce important statistical information on a given dataset and to identify unmasked statistical problems when investigating a given financial phenomenon.

We analyze three specific models used for the analysis of the short interest rate. The first one is a much quoted model proposed by [Chan et al. \(1992\)](#) (hereafter CKLS), that nests several linear-drift one-factor models for the short rate process. The main result of this paper is that the more appropriate models for the US short interest rates over the period 1964 to 1989 are those that allow the conditional volatility of short interest rate changes to

---

<sup>1</sup> See [Knez and Ready \(1997\)](#) for a recent exception.

be highly dependent on the level of the short rate. We choose this first model because it is a typical application of GMM in finance, where estimation and testing procedures are derived from highly nonlinear orthogonality functions, and because it has been criticized along many dimensions. On the one hand, there is a debate in the literature to understand whether we need sophisticated multi-factor models (Dai and Singleton, 2000). On the other hand, more complex single factor models such as regime-switching models,<sup>2</sup> models allowing for nonlinearities in the drift term,<sup>3</sup> or models adding GARCH and similar features<sup>4</sup> still constitute the workhorse of many recent term structure applications.

When the CKLS models are analyzed with the RGMM, they are all clearly rejected. The different results produced by the classical GMM are explained and interpreted by performing a simple sensitivity experiment, where we show that the classical GMM results are highly unstable. Specifically, we observe strong instabilities of the GMM test results with respect to relatively large changes in one influential observation and to very small changes of a larger set of no more than 10 influential observations identified by the RGMM.

Looking at the influential observations found by means of the RGMM, we identify those observations that are primarily responsible for the different results between the robust and the classical analysis. In fact, we note that the majority of them are clustered in the 1979–1982 period, a time span that is well known to coincide with a temporary change in the monetary policy of the Federal Reserve.<sup>5</sup> This anomalous clustering of influential points may suggest a change of structure over this period, rather than the existence of a set of isolated outliers. This confirms the results in the literature (e.g. switching models), where this period is found to be the main cause of misspecification in the CKLS setting. It is important to stress, however, that with the RGMM we obtain this same conclusion without changing the basic model but by slightly modifying the estimation and testing procedure.

To illustrate this important point further and to emphasize the chance to discover, using RGMM, a possible misspecification within the original model setting, we re-analyze the findings of two recent articles attempting to extend the CKLS framework in two particular ways. Among several possible specifications we choose two, which can be estimated with GMM and which are direct extensions of the CKLS models. In particular, as a first example, we analyze a model with a quadratic drift term used by Ahn and Gao (1999) on monthly data.<sup>6</sup> In this application, we find that while with the classical GMM the quadratic drift model is not rejected, with the RGMM it is. Here too, we find highly unstable GMM test statistics and a clustering of influential observations during the 1979–1982 period, indicating that adding a quadratic drift term in this case does not help to explain this particular period.

As a second simple extension of the CKLS models, we further analyze a model that allows for a structural shift in the parameters (see also CKLS, Bliss and Smith, 1999;

<sup>2</sup> See Cai (1994), Gray (1996), Ang and Bekaert (2002b), Bekaert et al. (2001).

<sup>3</sup> See Aït-Sahalia (1996), Stanton (1997), Jiang (1998) and Ahn and Gao (1999).

<sup>4</sup> See Brenner et al. (1996), Koedijk et al. (1997) and Ball and Torous (1999).

<sup>5</sup> Further influential points show patterns similar to those of the probabilities estimated in switching regime models (cf. for instance Gray, 1996) and in particular corresponding to the first OPEC crisis and the October 1987 Stock market crash.

<sup>6</sup> A similar specification was used by Aït-Sahalia (1996) in a semi-parametric setting on daily data.

Brenner et al., 1996). Bliss and Smith (1999) define a temporary parameter shift between 1979 and 1982, which seems a natural model extension based on both the evidence for a structural break during this sub-period in the literature and the robust weights obtained by the RGMM in the CKLS case. Also for this extension, it turns out that the classical GMM fails to reject the null hypothesis of no parameter shift while the RGMM rejects the unrestricted CKLS model in favor of the temporary structural break model. However, when we analyze the weights obtained for the models including a temporary parameter shift, we find that they are again clustered in the 1979–1982 period, with similar patterns as those identified earlier in the constant parameters models. We therefore conclude that a CKLS model with a temporary parameter shift is also unable to take adequately into account the Fed experiment.

Finally, we observe that the structure of the estimated robust weights suggests that a CIR model could be supported by the data after 1982, because in this time span only a few isolated influential points are found. This last finding is consistent with Zhang (1999), who would not reject a CIR specification for the US interest rate series in the post 1982 period. Similar conclusions arise for the Euro-mark case.

The remainder of the paper is organized as follows. Section 2 introduces some basic definitions, the RGMM methodology and the models under scrutiny. We present empirical results in Section 3 where we discuss the sensitivity analyses of the GMM specification test and the results of the RGMM estimation and model selection strategy. Section 4 concludes the paper with some summarizing remarks.

## 2. RGMM analysis of models for the short rate process

### 2.1. The CKLS framework and some extensions

CKLS consider a class of widely used stochastic processes for the short rate that are nested within the stochastic differential equation

$$dr_t = (\alpha + \beta r_t)dt + \sigma r_t^\gamma dW_t, \quad (1)$$

where  $r_t$  is the short rate at time  $t$  and  $(W_t)_{t \geq 0}$  is a standard Brownian motion in  $\mathbb{R}$ . The scalars  $\alpha$  and  $\beta$  characterize the linear drift component,  $\sigma$  is the instantaneous volatility parameter, while the constant  $\gamma$  measures the sensitivity of volatility with respect to the current interest rate level  $r_t$ .

By imposing a set of restrictions on  $(\alpha, \beta, \sigma, \gamma)$  we obtain a well-known class of models for the short rate. For completeness, these models are listed in Table 1 with the corresponding parameter restrictions.

A discretization of Eq. (1) yields the discrete-time model

$$r_t - r_{t-1} = \alpha + \beta r_{t-1} + \varepsilon_t, \quad (2)$$

where  $E(\varepsilon_t) = 0$  and  $E(\varepsilon_t^2) = \sigma^2 r_{t-1}^{2\gamma}$ .

Table 1  
Alternative models of the short rate<sup>a</sup>

| Model                           | $\alpha$ | $\beta$ | $\sigma$ | $\gamma$ | Restrictions                            |
|---------------------------------|----------|---------|----------|----------|-----------------------------------------|
| Merton                          |          | 0       |          | 0        | (0 attainable)                          |
| Vasicek                         |          |         |          | 0        | $\beta < 0$ (0 attainable)              |
| Cox–Ingersoll–Ross              |          |         |          | 1/2      | $\beta < 0$ and $2\alpha \geq \sigma^2$ |
| Dothan                          | 0        | 0       |          | 1        | –                                       |
| Geometric Brownian Motion       | 0        |         |          | 1        | $\beta < 0$ (0 attainable)              |
| Brennan Schwartz                |          |         |          | 1        | $\beta < 0$ and $\alpha > 0$            |
| Variable Rate                   | 0        | 0       |          | 3/2      | (0 attainable)                          |
| Constant Elasticity of Variance | 0        |         |          |          | $\beta < 0$ (0 attainable)              |

<sup>a</sup> Natural restrictions have to be imposed on the parameter values to ensure that the drift is mean-reverting at high interest rate values (infinity not attainable) and zero is unattainable; see [Ait-Sahalia \(1996\)](#).

The CKLS orthogonality conditions suited for a GMM estimation of Eq. (2) are

$$E(\varepsilon_t) = 0 \quad E(\varepsilon_t r_{t-1}) = 0 \quad E(\eta_t) = 0 \quad E(\eta_t r_{t-1}) = 0, \quad (3)$$

where  $\eta_t = \varepsilon_t^2 - \sigma^2 r_{t-1}^{2\gamma}$ .

The CKLS framework may be naturally extended in order to include a nonlinear drift. [Ahn and Gao \(1999\)](#), for example, estimate the parameters of an extension of Eq. (1) where the drift is quadratic and  $\gamma = 1.5$ ,<sup>7</sup>

$$dr_t = (\alpha_1 + \alpha_2 r_t + \alpha_3 r_t^2)dt + \sigma r_t^{1.5} dW_t, \quad (4)$$

using the orthogonality conditions:

$$\begin{aligned} E(\varepsilon_t) &= 0 \quad E(\varepsilon_t r_{t-1}) = 0 \quad E(\varepsilon_t r_{t-1}^2) = 0 \\ E(\eta_t) &= 0 \quad E(\eta_t r_{t-1}) = 0 \quad E(\eta_t r_{t-1}^3) = 0, \end{aligned} \quad (5)$$

where  $\varepsilon_t = r_t - r_{t-1} - (\alpha_1 + \alpha_2 r_{t-1} + \alpha_3 r_{t-1}^2)$  and  $\eta_t = \varepsilon_t^2 - \sigma^2 r_{t-1}^{2\gamma}$ .

A simple GMM testing procedure for a structural break as used in CKLS and in [Bliss and Smith \(1999\)](#) can be obtained by extending model (2) to

$$r_t - r_{t-1} = (\alpha + \delta_1 D_t) + (\beta + \delta_2 D_t) r_{t-1} + \varepsilon'_t, \quad (6)$$

where  $E(\varepsilon'_t) = 0$  and  $E(\varepsilon'_t{}^2) = (\sigma^2 + \delta_3 D_t) r_{t-1}^{2(\gamma + \delta_4 D_t)}$ , and  $D_t$  is a dummy variable. The orthogonality conditions used for a GMM estimation of this model are

$$\begin{aligned} E(\varepsilon'_t) &= 0 \quad E(\varepsilon'_t r_{t-1}) = 0 \quad E(\eta'_t) = 0 \quad E(\eta'_t r_{t-1}) = 0 \\ E(\varepsilon'_t D_t) &= 0 \quad E(\varepsilon'_t r_{t-1} D_t) = 0 \quad E(\eta'_t D_t) = 0 \quad E(\eta'_t r_{t-1} D_t) = 0, \end{aligned} \quad (7)$$

where  $\eta'_t = \varepsilon'_t{}^2 - (\sigma^2 + \delta_3 D_t) r_{t-1}^{2(\gamma + \delta_4 D_t)}$ .

<sup>7</sup> The conditions  $\alpha_3 \leq 0$  (or  $\alpha_2 < 0$  if  $\alpha_3 = 0$ ) and  $\alpha_1 \geq 0$  (or  $\alpha_2 > 0$  if  $\alpha_1 = 0$ ), ensure that infinity and zero are unattainable; cf. [Ahn and Gao \(1999, p. 731\)](#).

For example, CKLS test for a permanent parameter shift after October 1979 by setting a dummy equal to 0 before October 1979 and to 1 afterwards, while Bliss and Smith (1999) test for a temporary parameter shift from October 1979 through September 1982 by setting the dummy equal to 1 only from October 1979 to September 1982. As noted in the Introduction, the second specification is consistent with the widespread economic knowledge of a temporary structural break in the Fed's monetary policy (the so-called Fed experiment, which started in 1979 and ended in 1982) and with the robust weights obtained in the empirical analysis of the next section.

The sets of orthogonality conditions (3), (5) and (7) are defined by an unbounded orthogonality function and imply non-robust GMM estimators and tests (see Section 2.3 for a more detailed explanation of this point).

## 2.2. An illustrative example

In order to illustrate some of the effects of the contamination by outliers, we generate 200 paths of 300 observations from a discrete-time CIR process (Eq. (2)) with parameter values  $\alpha = 0.074$ ,  $\beta = -0.86$  and  $\sigma = 0.156$ . The simulated sample paths are constructed assuming that the length of time between observations of the diffusion is  $\Delta = 1/12$ . We choose these parameter values in order to produce a monthly autocorrelation and a variability comparable with the CKLS dataset. We then add an outlier comparable with the one identified by RGMM in March/April 1980 using the CKLS dataset (see Section 3.2). The average change in the interest rate for the generated series is about 0.5%. We perform a sensitivity analysis by taking the largest observation<sup>8</sup> and by contaminating it with 300 and 400 basis points, respectively. We then compute Hansen's test and its robust version at a 5% nominal level. The latter is Hansen's over-identification  $\chi^2$  test based on a robust GMM estimation (see Section 2.3). Note that the numerical value of the contaminated observation is not exaggerated when compared with short rates observed during the 1979–1982 period. In the CKLS dataset, for example, the variability of the realized short-term rate during this period is high, varying from 15% to 9.5% in March 1980.

We want to show the effects of contamination when comparing two models in the CKLS framework. As a particular alternative we choose the Brennan Schwartz model. Similar results are obtained with other models. Table 2 reports rejection rates for the CIR and the Brennan Schwartz hypothesis using the classical and the robust Hansen's test. We see that in the absence of outliers, the classical Hansen's test performs satisfactorily. First, as expected, the rejection rate of the Brennan Schwartz specification is higher than the one obtained using the robust Hansen's test. Second, the empirical size of the CIR hypothesis is 8%, but this can be due to sampling variability.

When the series is contaminated by an outlier, the rejection rate for the Brennan Schwartz model suddenly decreases in the first contamination case (300 basis points) to 51% and in the second contamination case (400 basis points) to 12%, while the empirical size of the CIR hypothesis drops to 0% in both cases. The robust test seems to perform acceptably in both cases.

<sup>8</sup> We have chosen to contaminate only the largest observation in order to provide an example close to the data structure in CKLS. Similar results arise when more observations are contaminated by smaller amounts.

Table 2  
An illustrative example

| Model            | Non-contaminated series | Contaminate by 300 basis points | Contaminate by 400 basis points |
|------------------|-------------------------|---------------------------------|---------------------------------|
| GMM: reject CIR  | 8                       | 0                               | 0                               |
| RGMM: reject CIR | 5                       | 5                               | 5                               |
| GMM: reject BS   | 92                      | 51                              | 12                              |
| RGMM: reject BS  | 80                      | 64                              | 60                              |

Empirical size and power of the GMM and RGMM when the underlying series of 300 observations is generated by a CIR model and when an outlier of 300 and 400 basis points, respectively, is added. The table reports the rejection rates in percentage when testing for a CIR and a Brennan Schwartz (BS) model with GMM and RGMM. The table values are the percentages of rejections of the specific model. The total number of runs was set to 200. The nominal size of the test is 5%. The tuning constant  $c$  for the robust test is set to 5.85.

From this simple example, we see that the classical test is not able to distinguish correctly between two competing models, once a series is contaminated. In particular, for the chosen type of contamination the models with high  $\gamma$  tend to be erroneously selected as in the CKLS case. In addition, we notice that this is similar to the results for the Euro-DM data (Section 3.4), where the classical GMM test fails to distinguish between different models while the robust GMM test selects the CIR specification clearly.

### 2.3. Robust GMM estimation and testing

This section describes briefly the RGMM methodology and outlines some of its features in a simplified CKLS setting. Full details are provided in [Ronchetti and Trojani \(2001\)](#). A simplified example is discussed in Section 2.4.

Let  $\{X_t\}_{t \in \mathbb{N}}$  be an ergodic series and  $\theta_0$  the unknown model parameter to be estimated. Further, let  $h: \mathbb{R}^N \times \Theta \rightarrow \mathbb{R}^H$  be an orthogonality function defining a set of orthogonality conditions

$$E_F h(X_1; \theta_0) = 0. \quad (8)$$

We stress the dependence of this orthogonality condition on the underlying model distribution  $F$ , since it is precisely the form of this functional relation that determines the robustness of a GMM statistic. Let us denote by  $W := (W_n)_{n \in \mathbb{N}}$  a set of weighting matrices converging a.s. to

$$W_0 := [E_F(h(X_1; \theta_0)h'(X_1; \theta_0))]^{-1}.$$

An optimal generalized method of moments estimator  $\tilde{\theta} := (\tilde{\theta}_n)_{n \in \mathbb{N}}$  of the model parameter  $\theta_0$  is defined as a sequence of solutions to the optimization problem (see [Hansen, 1982](#))

$$\min_{\theta \in \Theta} E_{F_n} h'(X_1; \theta) W_n E_{F_n} h(X_1; \theta),$$

where  $F_n := \frac{1}{n} \sum_{i=1}^n \delta_{X_i}$  is the empirical distribution of  $X_1, \dots, X_n$ , and  $\delta_x$  denotes the point mass distribution at  $x \in \mathbb{R}^N$ .

The influence function (IF)<sup>9</sup> of the GMM estimator is given by

$$\text{IF}(x; \tilde{\theta}, F) = - \left[ E_F \frac{\partial h'(X_1; \theta_0)}{\partial \theta} W_0 E_F \frac{\partial h(X_1; \theta_0)}{\partial \theta'} \right]^{-1} E_F \frac{\partial h'(X_1; \theta_0)}{\partial \theta} W_0 h(x; \theta_0). \quad (9)$$

The IF of a statistical functional describes the linearized asymptotic bias of a statistic under single point contaminations  $\delta_x$  of the assumed model distribution  $F$ . Moreover, under weak regularity conditions, boundedness of the IF is sufficient to control the bias of a statistic in an  $\varepsilon$ -neighborhood of the model  $F$  defined by

$$(1 - \varepsilon)F + \varepsilon G,$$

where  $G$  is an arbitrary contaminating distribution. An unbounded IF implies an unbounded bias of a statistic (an estimator, the power or the level of a test) already under small single point contaminations of the underlying model. In order to avoid unbounded bias when the model is slightly contaminated, a bounded IF is necessary. Notice (see Eq. (9)) that the IF of a GMM estimator is linearly related to  $h(\cdot; \theta_0)$ . Therefore, a GMM estimator has a bounded IF if and only if the function inducing the orthogonality conditions of the model is bounded in the observations.

The main idea behind the construction of a robust GMM estimator is to construct a weighted version of  $h$  that is bounded and that may again be interpreted as a set of GMM orthogonality conditions. For a given constant  $c > \sqrt{H}$  define a new mapping  $h_c^{A, \tau}: \mathbb{R}^N \times \Theta \rightarrow \mathbb{R}^H$  by

$$h_c^{A, \tau}(x, \theta) = A[h(x; \theta) - \tau]w_c[A[h(x; \theta) - \tau]], \quad (10)$$

where  $w_c(y) := \min(1, c/\|y\|)$  for  $y \neq 0$ , and  $w_c(0) := 1$ , and the nonsingular matrix  $A$  and the vector  $\tau$  are determined by the implicit equations:

$$E_F h_c^{A, \tau}(X_1, \theta_0) = 0 \quad (11)$$

and

$$E_{F_n} h_c^{A, \tau}(X_1, \theta_0) h_c^{A, \tau}(X_1, \theta_0)' = I. \quad (12)$$

The RGMM estimator  $\tilde{\theta}_c^{A, \tau}$  is obtained by iteratively computing the GMM estimator associated with the bounded orthogonality function  $h_c^{A, \tau}$ . Note that the expected value (Eq. (11)) is computed with respect to the given reference model distribution and not with respect to the empirical distribution of the data. In this paper, as a reference model for Eq. (2), we use one with conditionally normally distributed errors  $\varepsilon_i$ . Because of the robustness properties of the RGMM, the resulting estimator is stable in terms of bias and efficiency in

<sup>9</sup> Cf. Hampel (1974) and Hampel et al. (1986) for basic definitions and Ronchetti and Trojani (2001) for the GMM case.

a neighborhood of the reference model. The results are not too sensitive to slightly different choices of the reference model.<sup>10</sup>

The orthogonality function  $h_c^{A,\tau}$  is a truncated version of  $h$ . It takes into account the realistic case where only the majority of the data can reasonably fit the original moment conditions by re-weighting observations that are less compatible with the given GMM structure. The weights  $w_c(A[h(x;\theta) - \tau])$  assigned to each single observation  $x$  are used to detect outlying structures in the data.

Let us now briefly discuss how RGMM testing procedures are obtained. The goal of robust testing procedures is to control the maximal bias on the level and the power of a test that may arise because of a slight distributional misspecification of a null or an alternative hypothesis. The tuning constant  $c$  for the above RGMM estimator can be chosen so that the bias of the level and the power of all GMM tests remain within some pre-specified bounds, for a maximal amount of contamination a researcher may expect, given some prior information on the nature of the available data.

Specifically, when the underlying null distribution is locally contaminated as

$$P_{\varepsilon,n,G}^0 = \left(1 - \frac{\varepsilon}{\sqrt{n}}\right)F + \frac{\varepsilon}{\sqrt{n}}G,$$

where  $G$  is an arbitrary contaminating distribution, the maximal bias on the level of a GMM test is bounded by the inequality

$$\lim_{n \rightarrow \infty} |\alpha(P_{\varepsilon,n,G}^0) - \alpha_0| \leq \mu(\varepsilon c)^2 + o(\varepsilon^2), \quad (13)$$

where  $\mu = \frac{1-\alpha_0}{2} - \frac{1}{2}C_{r+2}(\eta_{1-\alpha_0})$ ,  $C_r(\cdot)$  is the cumulative distribution function of a non-central  $\chi^2(r)$  distribution with  $r$  degrees of freedom,  $\eta_{1-\alpha_0}$  is the  $1 - \alpha_0$  quantile of a  $\chi^2(r)$  distribution,  $\alpha_0 = \alpha(F)$  is the nominal level of the test at the model, while  $\alpha(P_{\varepsilon,n,G}^0)$  is the asymptotic level under contamination. Similar inequalities hold for the power of a RGMM test (see Table 1 in Ronchetti and Trojani, 2001).

In Section 3 of the paper we use (Eq. (13)) to choose  $c$  so that the maximal bias on the level of the RGMM tests is below 0.5% for model contaminations  $\varepsilon = 5\%$ .

#### 2.4. A characterization of robustness in a simplified CKLS framework

Within the CKLS framework, the IF of a GMM estimator  $\tilde{\theta}$  can be computed, given the defining orthogonality conditions

$$E_F h(r_t, r_{t-1}, \alpha, \beta, \sigma, \gamma) = 0.$$

For illustration purposes, we consider a one-dimensional orthogonality function  $g$  (this is the third orthogonality condition used in CKLS) given by

$$g(\sigma) := h(r_t, r_{t-1}, \alpha, \beta, \sigma, \gamma) := (r_t - r_{t-1} - \alpha - \beta r_{t-1})^2 - \sigma^2 r_{t-1}^{2\gamma} \quad (14)$$

<sup>10</sup> A specification with conditionally normal errors has been used by Nowman (1997) in a maximum likelihood framework.

and assume that under the given short rate model

$$dr_t = (\alpha + \beta r_t)dt + \sigma r_t^\gamma dW_t \quad (15)$$

the parameters  $\alpha$ ,  $\beta$  and  $\gamma$  are known.

For this case, the IF of the GMM estimator  $\tilde{\sigma}$  is easily obtained by means of Eq. (9) as

$$\begin{aligned} \text{IF}((r_t, r_{t-1}); \tilde{\sigma}, F) &= - \left[ E_F \frac{\partial g}{\partial \sigma}(\sigma) W_0 E_F \frac{\partial g}{\partial \sigma}(\sigma) \right]^{-1} W_0 E_F \frac{\partial g}{\partial \sigma}(\sigma) h(r_t, r_{t-1}, \alpha, \beta, \sigma, \gamma) \\ &= -E_F [2\sigma r_{t-1}^{2\gamma}]^{-1} h(r_t, r_{t-1}, \alpha, \beta, \sigma, \gamma), \end{aligned} \quad (16)$$

where  $F$  is the distribution of  $(r_t, r_{t-1})$  under model (15). As expected, the IF of  $\tilde{\sigma}$  is unbounded, since  $h$  is a polynomial in  $r_t$  and  $r_{t-1}$  whose order depends on  $\gamma$ .

For example, in the variable rate model ( $\alpha = \beta = 0, \gamma = 1.5$ ) one obtains

$$\text{IF}((r_t, r_{t-1}), \tilde{\sigma}, F) = \frac{-((r_t - r_{t-1})^2 - \sigma^2 r_{t-1}^3)}{2\sigma E_F(r_{t-1}^3)}.$$

Similar structures arise for other CKLS models. In particular, the IF is proportional to  $-h$ .

As an illustration, we present in Fig. 1 a plot of the function  $-h$  implied by Eq. (14) for the Brennan Schwartz model.

The surface plot of the IF is (mainly) negative for single point contaminations  $(r_t, r_{t-1})$  of the short rate process ranging between 0 and 0.20. Specifically the negative bias of  $\tilde{\sigma}$  is close to zero for contaminations where  $r_t$  and  $r_{t-1}$  are approximately equal, that is when short rates do not vary excessively. On the other side, a serious negative bias arises when the contaminations are such that the probability of a large difference  $r_t - r_{t-1}$  is high, that is when short rates move very fast. One would thus expect a GMM estimate  $\tilde{\sigma}$  induced by functions  $h$  of this form not to be robust to such large interest rates changes. An even more extreme surface arises, if we look at the second graph in Fig. 1, where the sixth orthogonality condition used by Ahn and Gao (1999, p. 732), that is

$$E((\epsilon_t^2 - \sigma^2 r_{t-1}^3) r_{t-1}^3) = 0,$$

is plotted, which leads to an orthogonality function given by a polynomial of degree 6 in  $r_{t-1}$ .

A robust estimation of the parameters of a short rate model is only a part of a robust selection of models for the short rate. An even more important task is the construction of robust testing procedures for comparing different sub-models. We illustrate this second robustness issue within the simplified GMM setting (Eq. (14)).

Suppose we would like to test the GMM specification null hypothesis

$$\mathcal{H}_0 : E_F h(r_t, r_{t-1}; \theta_{\mathcal{H}_0}) = 0; \quad \theta_{\mathcal{H}_0} := (\alpha, \beta, \sigma_0, \gamma), \quad (17)$$

in model (15) using a GMM specification test. The statistical functional  $\xi$  determining the GMM specification test in our simplified setting is given by the quadratic form

$$\xi(F_n) := U^2(F_n) := W_0 (E_{F_n} h(r_t, r_{t-1}; \theta_{\mathcal{H}_0}))^2.$$

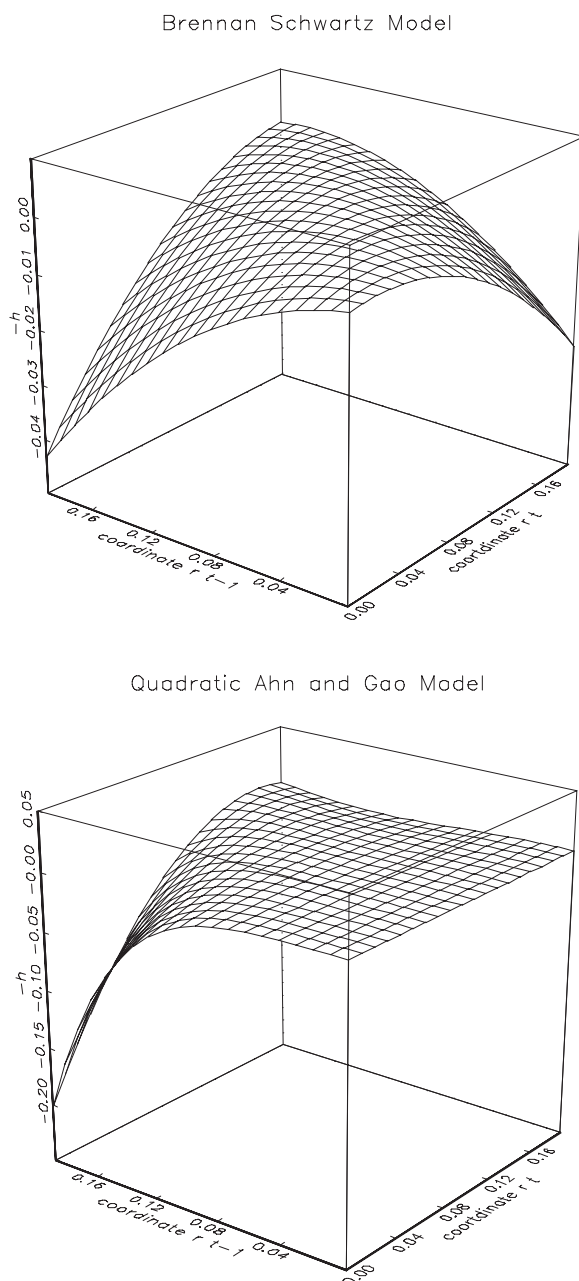


Fig. 1. Function  $-h$  (cf. Eq. (16)) in the Brennan Schwartz model and in the quadratic drift model of Ahn and Gao (1999) when only the parameter  $\sigma$  is estimated with one single orthogonality condition.

In the present example,  $n\zeta(F_n)$  is asymptotically  $\chi^2(1)$  distributed under  $\mathcal{H}_0$ , when Eq. (17) is satisfied by the data. Given a nominal level  $\alpha_0$ , a test of  $\mathcal{H}_0$  is then defined by a critical region given by

$$\{n\zeta(F_n) \geq \chi_{1-\alpha_0}^2(1)\},$$

where  $\chi_{1-\alpha_0}^2(1)$  is the  $1 - \alpha_0$  quantile of a  $\chi^2$  distribution with 1 *df*. However, when the distribution  $G$  of the data slightly differs from  $F$  this test can have a nominal level different from  $\alpha_0$ .

As in the estimation case, the distortion of the level that is obtained under single point contaminations of  $F$  can be approximated by the IF of a corresponding statistic. In this simple GMM setting, this bias is proportional to the squared IF of  $U$  (see Heritier and Ronchetti, 1994; Ronchetti and Trojani, 2001), i.e. in our case to  $h^2(r_t, r_{t-1}, \alpha, \beta, \sigma_0, \gamma)$ . Hence, one would expect the level of such a GMM test to be highly sensitive to large  $r_t - r_{t-1}$  observations. This is confirmed by the illustrative example of Section 2.2 and by simulations in Ronchetti and Trojani (2001) for a simple ARCH specification.

### 3. Data and empirical results

#### 3.1. Data

We first use the CKLS dataset, that is the 1-month Treasury Bill series taken from the 12-month Fama Treasury Bill Files included in the CRSP monthly Government Bonds Files. These are 307 monthly observations from June 1964 to December 1989. As pointed out by Duffee (1996), this data series may not be ideal for the type of analysis addressed here because of an idiosyncratic variation in the CKLS dataset when compared to other 1-month US rates series. Given the methodological focus of this paper, we first present our analysis with the CKLS dataset to guarantee comparability with the original CKLS work.<sup>11</sup>

In a second step, we use the McCulloch and Kwon (1993) dataset over the period from December 1946 to February 1991 for comparability with the Ahn and Gao (1999) study, which extends the CKLS model to a diffusion process with an additional quadratic drift term. This analysis is performed in order to show that the results are not specific to the CKLS model tested on the CKLS dataset and that similar findings arise when using one of the most recent CKLS extensions estimated with the GMM.

In a third step, we also test the RGMM methodology using the 1-month Euro-mark rates (taken from Datastream) over the period from February 1975 to April 2000, for a total of 303 observations. We perform this analysis on the Euro-mark essentially because we do not want our empirical assessment of the RGMM to be dependent on the very particular and largely studied structure of US short interest rates over the CKLS sample period.

<sup>11</sup> A further analysis over the 1975–1999 period using Eurodollar rates yields the same conclusions as those obtained using the CKLS sample.

Table 3  
Summary statistics

| Dataset  | Variables    | <i>N</i> | Mean    | SD      | $\rho_1$ | $\rho_2$ | $\rho_3$ | $\rho_4$ | $\rho_5$ |
|----------|--------------|----------|---------|---------|----------|----------|----------|----------|----------|
| CKLS     | $r_t$        | 307      | 0.06715 | 0.02675 | 0.95     | 0.91     | 0.86     | 0.82     | 0.8      |
|          | $\Delta r_t$ | 306      | 0.00009 | 0.00821 | −0.08    | 0.07     | −0.12    | −0.14    | −0.03    |
| AG       | $r_t$        | 531      | 0.0482  | 0.0319  | 0.98     | 0.96     | 0.94     | 0.93     | 0.91     |
|          | $\Delta r_t$ | 530      | 0.0001  | 0.0061  | 0.022    | −0.02    | −0.117   | −0.077   | −0.017   |
| Euromark | $r_t$        | 303      | 0.0574  | 0.02459 | 0.981    | 0.964    | 0.948    | 0.928    | 0.904    |
|          | $\Delta r_t$ | 302      | 0.0001  | 0.0046  | −0.049   | −0.020   | 0.145    | 0.093    | 0.112    |

Means, standard deviations, and autocorrelations of monthly interest rates from June 1964 to December 1989 for the CKLS dataset, from December 1946 to February 1991 for the McCulloch and Kwon dataset, and from February 1975 to April 2000 for the Euro-mark dataset.  $r_t$  denotes the short rate at time  $t$ , while  $\Delta r_t = r_{t+1} - r_t$  is the associated monthly interest rate change.  $\rho_j$  denotes the  $j$ th-order autocorrelation coefficient.

Table 3 reports the means, standard deviations, and first five autocorrelations of the 1-month interest rate for the CKLS, the AG and the Euro-mark dataset.

### 3.2. RGMM estimation and testing of the CKLS models

We estimate the CKLS models (Eq. (2)) with the RGMM and compare the parameter estimates with their classical counterparts. The left-hand side of Table 4 shows the classical GMM estimates and Hansen's statistics. This table corresponds to Table 3 in CKLS.<sup>12</sup>

As in CKLS, models that allow for values  $\gamma \geq 1$  are not rejected by Hansen's statistic, while models where  $\gamma \in [0, 1)$  are. The RGMM estimates are presented on the right-hand side of Table 4, where the bound on the RGMM estimator for the unrestricted CKLS model was set to  $c = 5.85$ .<sup>13</sup>

With this bound, we identify 14 observations out of 307 as influential with corresponding weights  $w_c(\cdot)$  defined in Eq. (10) that were less than 1. Using the bounded orthogonality function (10) the constrained models are estimated. Based on the robust Hansen's statistics, all models with  $\gamma \leq 1$  are rejected at a 5% significance level. This is at odds with the classical result, where models with  $\gamma = 1$  could not be rejected. Moreover, the significance of rejection of the single models is high, and the  $p$ -value of Hansen's test changes when comparing the classical with the robust GMM tests. For the Geometric Brownian Motion model, for example, the  $p$ -value goes from about 0.2 to about 0.00003. Models with  $\gamma > 1$  are not rejected by standard significance levels.<sup>14</sup> Note that the results for  $\gamma > 1$  are given only for the sake of completeness and comparison with CKLS, since the standard GMM asymptotic inference does not hold under this parameter choice (see Broze et al., 1995).

<sup>12</sup> Compared to CKLS, our data have been taken to be the annualized interest rate. Notice that the relevant  $\gamma$  estimates and test results are invariant to scaling.

<sup>13</sup> This corresponds to a maximal bias of 0.5% in the asymptotic size of the test for a model contamination of 5%; cf. Section 2.3. Different choices of  $c$  give qualitatively similar results.

<sup>14</sup> Moreover, a detailed inspection of the influential points suggests that the unrestricted CKLS model is also misspecified; see below.

Table 4

Classical and robust GMM estimates of alternative models for the short-term interest rate

| Model        | Classical GMM    |                    |                   |                  |                 | Robust GMM       |                    |                   |                   |                  | df |
|--------------|------------------|--------------------|-------------------|------------------|-----------------|------------------|--------------------|-------------------|-------------------|------------------|----|
|              | $\alpha$         | $\beta$            | $\sigma$          | $\gamma$         | $\xi$           | $\alpha$         | $\beta$            | $\sigma$          | $\gamma$          | $\xi$            |    |
| Unrestricted | 0.0034<br>(1.85) | −0.0493<br>(−1.55) | 0.3656<br>(1.56)  | 1.4999<br>(5.95) | —               | 0.0024<br>(1.35) | −0.0306<br>(−1.02) | 0.3843<br>(2.52)  | 1.5371<br>(9.59)  | —                | —  |
| Merton       | 0.0005<br>(1.44) | 0                  | 0.0062<br>(14.54) | 0                | 6.76<br>(0.034) | 0.0005<br>(1.48) | 0                  | 0.0052<br>(14.70) | 0                 | 43.88<br>(0.000) | 2  |
| Vasicek      | 0.0005<br>(0.33) | −0.0013<br>(−0.04) | 0.0062<br>(14.33) | 0                | 6.80<br>(0.009) | 0.0023<br>(2.14) | −0.0293<br>(−1.63) | 0.0048<br>(17.32) | 0                 | 65.71<br>(0.000) | 1  |
| CIR          | 0.0011<br>(0.67) | −0.0102<br>(−0.36) | 0.0254<br>(15.28) | 1/2              | 4.90<br>(0.027) | 0.0025<br>(2.38) | −0.0330<br>(−1.83) | 0.0215<br>(20.34) | 1/2               | 53.84<br>(0.000) | 1  |
| Dothan       | 0                | 0                  | 0.0320<br>(15.94) | 1                | 5.60<br>(0.133) | 0                | 0                  | 0.0954<br>(18.57) | 1                 | 19.83<br>(0.000) | 3  |
| GBM          | 0                | 0.0084<br>(1.50)   | 0.0993<br>(16.07) | 1                | 3.16<br>(0.206) | 0                | 0.0070<br>(1.34)   | 0.0951<br>(20.10) | 1                 | 20.66<br>(0.000) | 2  |
| BS           | 0.0020<br>(1.24) | −0.0262<br>(−0.92) | 0.0994<br>(16.18) | 1                | 2.21<br>(0.137) | 0.0025<br>(2.40) | −0.0330<br>(−1.84) | 0.0914<br>(23.88) | 1                 | 29.27<br>(0.000) | 1  |
| VR           | 0                | 0                  | 0.05<br>(15.66)   | 3/2              | 6.31<br>(0.098) | 0                | 0                  | 0.3485<br>(19.79) | 3/2               | 5.39<br>(0.145)  | 3  |
| CEV          | 0                | 0.0086<br>(1.53)   | 0.1554<br>(1.18)  | 1.1711<br>(3.59) | 2.98<br>(0.084) | 0                | 0.0073<br>(1.54)   | 0.4619<br>(3.50)  | 1.6244<br>(13.97) | 2.88<br>(0.08)   | 1  |

The parameters are estimated by the classical GMM induced by the original orthogonality function  $h$  and by means of the RGMM induced by the modified orthogonality function  $h_{c,\tau}^{t,\tau}$  implied by a bound  $c=5.85$  on the influence function for the unrestricted model;  $t$ -statistics are in parentheses. The value of Hansen's statistics ( $\xi$  for short) are reported with  $p$ -values in parentheses and associated degrees of freedom ( $df$ ). The parameters are estimated from the following discrete-time system of equations:

$$r_t - r_{t-1} = \alpha + \beta r_{t-1} + \varepsilon_t,$$

$$E(\varepsilon_t) = 0, \quad E(\varepsilon_t^2) = \sigma^2 r_{t-1}^\gamma.$$

In order to verify the degree of data contamination, we take a closer look at the identified influential observations in Fig. 2 (Panels A and B), where the weights produced by the RGMM estimator are plotted together with the corresponding series for the short rate.

We identify automatically 14 influential observations through the robust weights. Ten of them occur in the 1979–1982 period. Consistently with the results of other studies (see Bliss and Smith, 1999; Gray, 1996; Brenner et al., 1996), we interpret this as a hint of a model misspecification of the CKLS models, caused by the change in the monetary policy of the Federal Reserve. Further inspection shows patterns of influential observations that are similar to regimes of high volatility and of high mean reversion found in Gray (1996) and Ball and Torous (1995). These influential points correspond to the first oil crisis and to the October 1987 stock market crash.

### 3.2.1. Robustness of the CKLS findings

To understand the robustness problem of the classical GMM specification test as well as the difference between a cross-validation technique for the treatment of single outliers (as used for instance in Bliss and Smith, 1999) and the RGMM, we perform some sensitivity

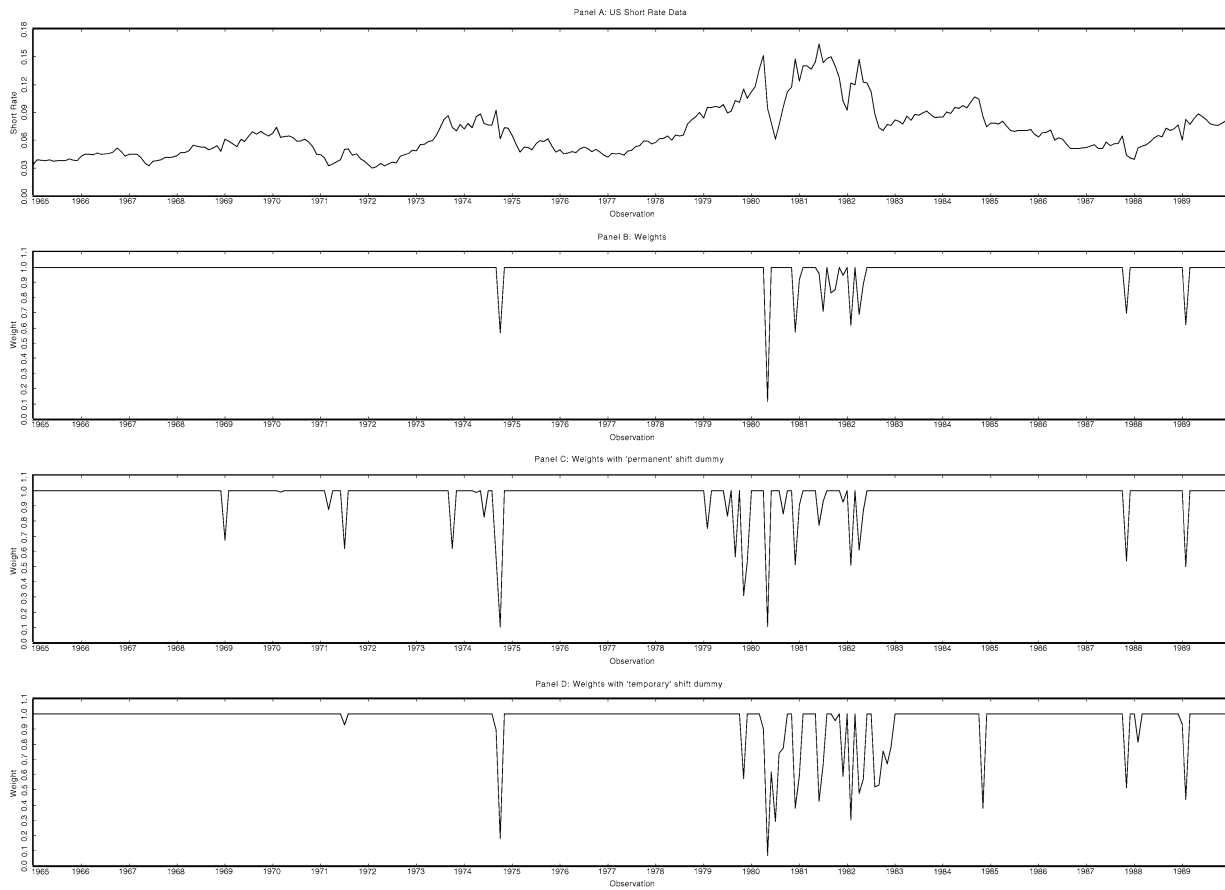


Fig. 2. USD Short Term Rate (Panel A) and weights of the RGMM Estimation (Panel B), of the ‘permanent’ dummy setting (Panel C) and of the ‘temporary’ dummy setting (Panel D). Influential observations correspond to weights that are smaller than 1.

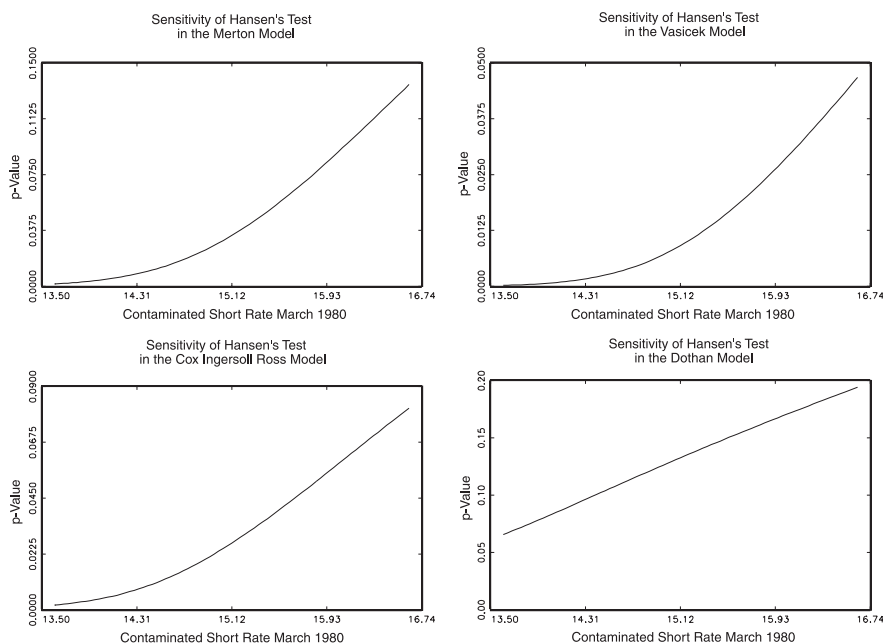


Fig. 3. Sensitivity analysis for the classical GMM-Specification Test in the Merton, the Vasicek, the Cox, Ingersoll, Ross and the Dothan Model. The March 1980 observation was contaminated.

analyses of the  $p$ -value of Hansen's statistic under different kinds of model contaminations.

We begin by investigating the robustness of Hansen's test with respect to single point contaminations. This corresponds to a contamination  $\varepsilon=1/306$ , that is approximately around 0.3% of the data. In particular, we perform this first analysis with respect to changes of the most influential observation corresponding to March 1980. The observed short-term rate is 0.1512. We vary this value between 0.1350 and 0.1674 by steps of size 0.001 and each time we recompute the  $p$ -value of Hansen's test for the nested models presented in Table 1.<sup>15</sup>

A general inspection of Fig. 3 shows steep  $p$ -values curves. Therefore, a small change of the short-term rate on the  $x$ -axis implies a large change of the  $p$ -values on the  $y$ -axis (the value in the middle of the  $x$ -axis represents the observed value of the short-term rate, i.e. 0.1512).

For the Merton, the CIR and the Variable Rate models, a change of the short rate level from 40 to 90 basis points is enough to obtain  $p$ -values not rejecting the model specification at a 5% significance level in the first two cases and rejecting it in the third. In the Geometric Brownian Motion, the Brennan Schwartz, the Dothan and the Constant

<sup>15</sup> The variability of the realized short-term rate changes around March 1980 is high, with a change from a 15% to a 9.5% interest rate level just after March 1980. Therefore, the magnitude of the contamination intervals in the sensitivity analysis is not unrealistic with respect to the structure of the short-rate observations over this particular period.

Elasticity of Variance models (not shown here), changes of 160 basis points are not enough to obtain reversions of the test decision. However, the variability of the corresponding  $p$ -values is high (ranging between 0.09 and 0.31 in the first case, between 0.09 and 0.18 in the second, between 0.06 and 0.20 in the third and between 0.05 and 0.13 in the fourth). Moreover, for larger contaminations a reversion of the classical test decision occurs, even with respect to these models. The same sensitivity analysis of Hansen's test derived from the RGMM estimator produces  $p$ -values that are very stable for all models. Moreover, the  $p$ -values for the first six nested models are smaller than  $10^{-3}$  for values of the short-term rate ranging between 0.1350 and 0.1674. For the last two models, the  $p$ -values are about 0.05–0.10.

The sensitivity of the classical GMM specification tests in this application is not limited to large contaminations of single observations. To illustrate this important point, we perform a further sensitivity analysis of the  $p$ -values of Hansen's test with respect to more general contaminations than those considered above. We move the six most influential observations (a contamination of about  $\varepsilon = 2\%$ ) over a grid contained in intervals of  $\pm 45$  basis points around the 6 observed non-contaminated values of the short rates. The maximal and the minimal  $p$ -values obtained for the respective CKLS models are listed in Table 5.

The maximal sensitivity of the  $p$ -values of Hansen's statistic is again comparable to that obtained for the above single points contaminations with  $p$ -values in the Merton, the Vasicek and the Variable Rate models, which in some cases reverse the test decision. In other models, a high sensitivity arises. In particular, in the Brennan Schwartz model, a clearly higher  $p$ -value sensitivity than in the single point contamination case is obtained (for this model the  $p$ -values are now changing from 0.0497 to 0.2338). Similar sensitivities in the  $p$ -values are obtained when moving the 10 most influential observations by  $\pm 20$  basis points. This confirms that when testing contaminated models by GMM even small changes of a few observations can have a strong impact on the test results. Notice, that the  $p$ -values of the robust GMM are stable and below  $10^{-3}$  for all models except the VR and the CEV, and are therefore not plotted.

Table 5

Maximal sensitivities of the  $p$ -values of Hansen's test under contamination of the six most influential observations

| Model                           | Minimal $p$ -value | Maximal $p$ -value |
|---------------------------------|--------------------|--------------------|
| Merton                          | 0.0082             | 0.1017             |
| Vasicek                         | 0.0016             | 0.0325             |
| Cox–Ingersoll–Ross              | 0.0059             | 0.0734             |
| Dothan                          | 0.0726             | 0.1811             |
| Geometric Brownian Motion       | 0.1211             | 0.2593             |
| Brennan Schwartz                | 0.0497             | 0.2338             |
| Variable Rate                   | 0.0392             | 0.1688             |
| Constant Elasticity of Variance | 0.0645             | 0.1072             |

The observations corresponding to August 1974, March 1980, November 1980, May 1981, October 1981 and March 1982 are contaminated in the CKLS dataset. The six observations are moved over a grid contained in intervals of  $\pm 45$  basis points around the original observation. The table reports the minimal and the maximal  $p$ -values for the CKLS models.

### 3.3. Extensions of the CKLS model

The analysis presented in Section 3.2 confirms the robustness problem of the CKLS model, particularly when the data sample spans the 1979–1982 period. Stanton (1997) and Aït-Sahalia (1996) have recently argued that single factor models with a nonlinear drift term can possibly cope with the high volatility, high mean-reverting 1979–1982 period. Alternatively, Ang and Bekaert (2002a) find that the nonlinearity of the drift function may be explained by means of a regime-switching model. Bliss and Smith (1999) and Brenner et al. (1996) allow a parameter shift in the 1979–1982 period and test for a possible structural break. Bliss and Smith (1999) find that when allowing the parameter to change over the 1979–1982 period, a CIR specification cannot be rejected. In this section we show that these recent extensions of the CKLS model face similar problems. Specifically, we extend our robust analysis to a model with a nonlinear drift term (the one used in Ahn and Gao, 1999), and to a model with a temporary dummy between October 1979 and September 1982 (as used in Bliss and Smith, 1999). Our main conclusion is that the stability problems and the clustering of the influential observations around the 1979–1982 period remain, even for these extensions.

#### 3.3.1. A short rate model with a nonlinear drift term<sup>16</sup>

In this section, we show that nonlinear drift extensions of a one-factor model may have potentially the same robustness problems described in the CKLS framework. We choose to replicate the Ahn and Gao (1999) study, i.e. specification (4), because it uses data at a monthly frequency as in CKLS and because it is estimated using GMM in its original setting. This specification is a special case of CKLS, when  $\alpha_3$  is set to 0. Ahn and Gao (1999) choose a  $\gamma$  value 1.5 in order to be broadly consistent with the values and shapes of the diffusion function found in CKLS and with the nonparametric and semiparametric analysis in Stanton (1997) and Aït-Sahalia (1996). Given the evidence in the preceding subsections, we should be careful when using  $\gamma = 1.5$ . However, since the aim of this study is mainly methodological, we adopt exactly the Ahn and Gao (1999) parameterization. Similar results arise for different values of  $\gamma$ . In the same way, we take only one of the many specification tests performed by Ahn and Gao (1999), since the main message remains the same.

The results for the classical and robust GMM estimation using the orthogonality conditions (Eq. (5)) are presented in Table 6.<sup>17</sup> First, we remark that the classical Hansen's test does not reject the Ahn and Gao (1999) model, while the robust test clearly does.

Second, as in the preceding section, the most informative contribution is provided by the estimated RGMM weights. Fig. 4 shows the influential points given by the RGMM

<sup>16</sup> The analysis of the quadratic short rate model of Ahn and Gao (1999) was suggested by an anonymous referee.

<sup>17</sup> Ahn and Gao (1999) estimate their models for two different sample sizes. The full sample goes from July 1946 to December 1991, the shorter sample from December 1960 to December 1991. We present the estimated parameters for the full sample size. Even more pronounced results arise using the shorter sample size.

Table 6  
Classical and RGMM estimates of the quadratic drift model

| Model        | $\alpha_1$         | $\alpha_2$       | $\alpha_3$         | $\sigma$          | $\xi$             | $df$ |
|--------------|--------------------|------------------|--------------------|-------------------|-------------------|------|
| Classical AG | –0.0697<br>(–1.08) | 0.0650<br>(1.73) | –0.0070<br>(–1.81) | 0.0370<br>(19.93) | 4.541<br>(0.103)  | 2    |
| Robust AG    | –0.0276<br>(–0.60) | 0.0434<br>(1.50) | –0.0050<br>(–1.50) | 0.0375<br>(26.17) | 21.276<br>(0.000) | 2    |

The parameters are estimated by means of the classical GMM induced by the original orthogonality function  $h$ —in the first row of the table—and by means of the RGMM estimator induced by the modified original orthogonality function  $h_c^{A,\tau}$  ( $c=6.58$ )—in the second row of the table—for the Ahn and Gao (1999) model (quadratic Ahn Gao). The  $t$ -statistics are in parentheses. The values of Hansen's statistics ( $\xi$  for brevity) are reported with  $p$ -values in parentheses and associated degrees of freedom ( $df$ ). The dataset used is the same as in Ahn and Gao (1999), spanning the period 12:1942 to 2:1991. The parameters are estimated from the following discrete-time system of equations:

$$r_t - r_{t-1} = \alpha_1 + \alpha_2 r_{t-1} + \alpha_3 r_{t-1}^2 + \varepsilon_t$$

$$E(\varepsilon_t) = 0, \quad E(\varepsilon_t^2) = \sigma^2 r_{t-1}^3.$$

when applied to the AG dataset. We notice that the pattern of the influential observations is similar to the CKLS case. In particular, there is again a clustering of influential observations in the 1979–1982 period. This suggests that the extension of Ahn and Gao (1999) is possibly still misspecified and in particular is not able to cope with the 1979–1982 period.

As in the CKLS case, the sensitivity of the classical GMM test statistics can be shown by contaminating the most influential observations. The results are not presented here for brevity's sake. Because of the higher order of the polynomials in the Ahn and Gao (1999) orthogonality functions, the overall shapes are even worse than those for the CKLS case of Section 3.2.1.<sup>18</sup>

### 3.3.2. A robust test for regime shift

A further straightforward extension used in CKLS, Bliss and Smith (1999) and Brenner et al. (1996) is to test for a possible break by allowing the parameters to change during given particular periods.

The classical and robust GMM results of model (6) for the temporary parameter shift specification are presented in Table 7. In this analysis, the CKLS dataset has been used in order to be comparable with the analysis in Bliss and Smith (1999).

The classical GMM test does not reject the hypothesis of parameter stability at a 5% significance level. By contrast, the RGMM statistic now rejects the GMM specification of the unrestricted CKLS model in favor of one allowing a temporary parameter shift between October 1979 and September 1982.<sup>19</sup>

<sup>18</sup> The results of this sensitivity analysis are available from the authors on request.

<sup>19</sup> Notice that the robust GMM test is tailored to reflect the structure of the majority of the data. Hence, it can be interpreted as a test of a flexible structural break hypothesis where the distribution implied by the change in monetary policy revealed by the data can deviate a little from that given in the dummy definition.

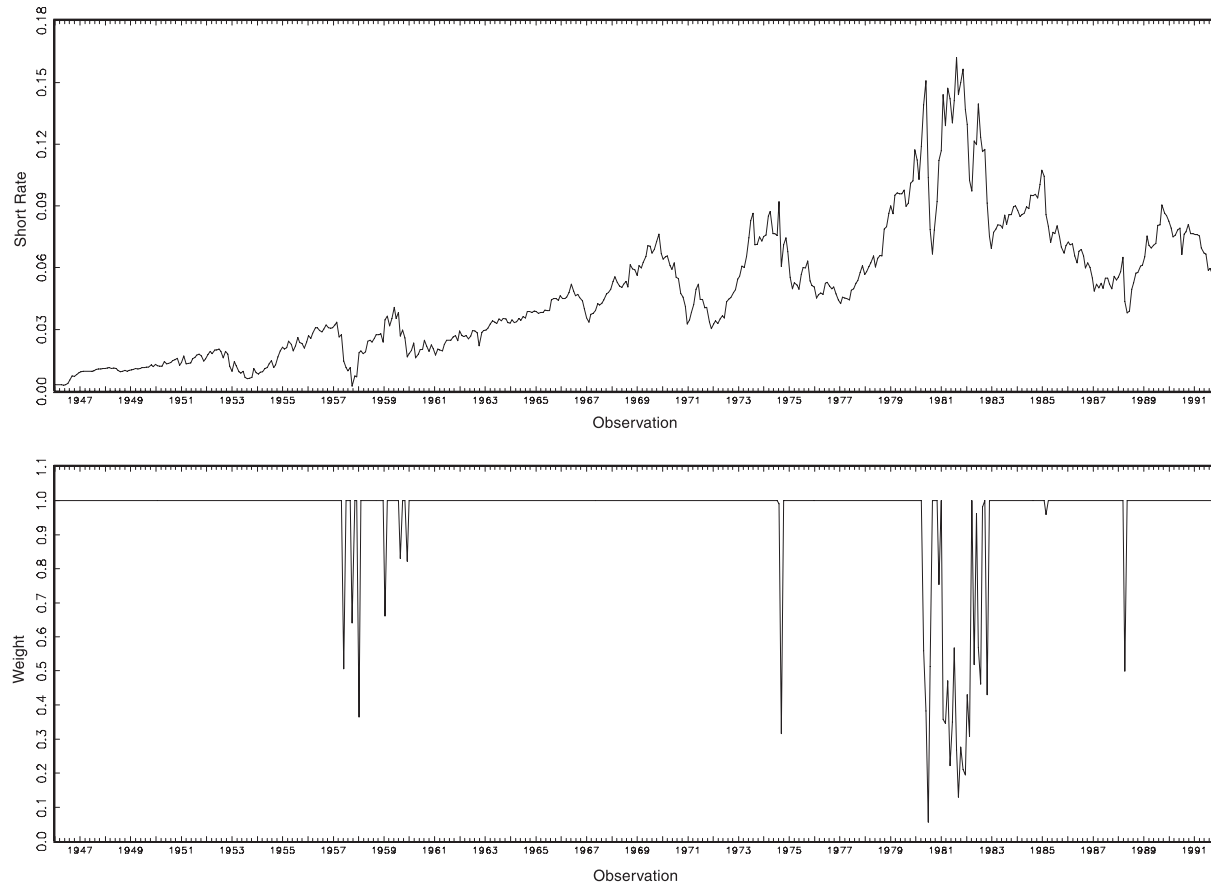


Fig. 4. Ahn and Gao (1999) data and weights implied by the RGMM estimator. Influential observations correspond to weights that are smaller than 1.

Table 7

Classical and robust GMM estimates of the unrestricted CKLS model with regime shift from October 1979 to September 1982

| $\alpha$           | $\beta$              | $\sigma$           | $\gamma$           | $\delta_1$         | $\delta_2$           | $\delta_3$           | $\delta_4$           | $\xi$             | $df$ |
|--------------------|----------------------|--------------------|--------------------|--------------------|----------------------|----------------------|----------------------|-------------------|------|
| <i>GMM</i>         |                      |                    |                    |                    |                      |                      |                      |                   |      |
| 0.002796<br>(2.44) | -0.043283<br>(-2.04) | 0.007794<br>(0.64) | 1.004215<br>(3.63) | 0.022736<br>(1.83) | -0.179365<br>(-1.64) | -0.005612<br>(-0.43) | -0.540319<br>(-0.95) | 8.265<br>(0.082)  | 4    |
| <i>RGMM</i>        |                      |                    |                    |                    |                      |                      |                      |                   |      |
| 0.001716<br>(1.65) | -0.021298<br>(-1.17) | 0.002287<br>(1.00) | 0.809461<br>(4.48) | 0.016947<br>(1.03) | -0.126808<br>(-0.95) | -0.001999<br>(-0.85) | -0.719627<br>(-1.45) | 10.358<br>(0.035) | 4    |

The parameters are estimated by means of the classical GMM induced by the original orthogonality function  $h$ —in the first row of the table—and by means of the RGMM estimator induced by the modified original orthogonality function  $h_c^{t,\tau}$  ( $c=6.5$ )—in the second row of the table—for the CKLS models with regime shift defined from October 1979 to September 1982;  $t$ -statistics are in parentheses. The value of Hansen's statistics ( $\xi$  for short) are reported with  $p$ -values in parentheses and associated degrees of freedom ( $df$ ). The parameters are estimated from the following discrete-time system of equations:

$$r_t - r_{t-1} = (\alpha + \delta_1 D_t) + (\beta + \delta_2 D_t) r_{t-1} + e_t'$$

$$E(e_t') = 0, \quad E(e_t'^2) = (\sigma^2 + \delta_3 D_t) r_{t-1}^{2(\gamma + \delta_4 D_t)}.$$

However, in Fig. 2 (Panel D) we observe again RGMM weights clustering around the 1979–1982 period, with similar patterns as before.<sup>20</sup> Hence, even with an ‘economically founded’ dummy a similar model misspecification as before persists, indicating that a CKLS model with deterministic parameter shift cannot provide a full description of the features of the CKLS dataset and more specifically the Fed Experiment.

### 3.4. Further evidence with Euro-mark data

Having verified the ability of the RGMM to identify well-known misspecification structures in the CKLS dataset, we apply the RGMM to a test of the CKLS models for the Euro-mark interest rates. For this series, the kind of misspecification that one could expect a priori in the data is less obvious and has been less investigated than in the US case where the Fed Experiment has clearly dominated the discussion.

The results obtained for the classical and robust GMM are presented in Table 8. When testing with the classical GMM, no model specification is rejected (with the exception of Merton's model) while only one model, the Cox–Ingersoll–Ross one, is not rejected by the RGMM<sup>21</sup> at a 5% confidence level.

Notice the striking difference with the classical GMM results, where essentially no model specification could be rejected. This is very similar to the lack of power demonstrated in the example of Section 2.2, where after adding an outlier to a series of

<sup>20</sup> Similar patches of outliers arise when testing for the (less natural) permanent structural break hypothesis for October 1979 onward, as in CKLS. The corresponding weights obtained by the robust GMM are shown in Fig. 2 (Panel C).

<sup>21</sup> As in Section 3.2, the robustness bound was set to  $c=5.85$ .

Table 8

Classical and robust GMM estimates of alternative models for the short-term Euro-DM interest rate

| Model        | Classical GMM    |                    |                     |                  |                 | Robust GMM         |                    |                     |                  |                  | df |
|--------------|------------------|--------------------|---------------------|------------------|-----------------|--------------------|--------------------|---------------------|------------------|------------------|----|
|              | $\alpha$         | $\beta$            | $\sigma$            | $\gamma$         | $\xi$           | $\alpha$           | $\beta$            | $\sigma$            | $\gamma$         | $\xi$            |    |
| Unrestricted | 0.1010<br>(1.24) | −0.0165<br>(−1.21) | 0.0997<br>(1.85)    | 0.8440<br>(2.43) | —               | 0.0966<br>(2.16)   | −0.0204<br>(−2.44) | −0.1669<br>(−4.49)  | 0.4033<br>(3.26) | —                | —  |
| Merton       | 0.0248<br>(0.56) | 0                  | 0.6874<br>(6.37)    | 0                | 6.96<br>(0.031) | −0.0039<br>(−0.22) | 0                  | 0.3084<br>(16.08)   | 0                | 11.66<br>(0.003) | 2  |
| Vasicek      | 0.1278<br>(2.23) | −0.0278<br>(−2.62) | 0.3341<br>(9.47)    | —                | 1.84<br>(0.176) | 0.0843<br>(2.07)   | −0.017<br>(−2.35)  | 0.3103<br>(16.35)   | 0                | 8.86<br>(0.003)  | 1  |
| CIR          | 0.1071<br>(1.82) | −0.0267<br>(−2.07) | 0.1654<br>(9.80)    | 1/2              | 0.60<br>(0.441) | 0.0976<br>(2.39)   | −0.0209<br>(−2.80) | 0.1402<br>(17.54)   | 1/2              | 0.49<br>(0.485)  | 1  |
| Dothan       | 0                | 0                  | 0.0815<br>(11.72)   | 1                | 1.45<br>(0.694) | 0                  | 0                  | 0.0496<br>(14.96)   | 1                | 20.76<br>(0.000) | 3  |
| GBM          | 0                | −0.0004<br>(−0.11) | 0.0814<br>(11.55)   | 1                | 1.44<br>(0.486) | 0                  | −0.0050<br>(−1.57) | 0.0493<br>(14.92)   | 1                | 19.26<br>(0.000) | 2  |
| BS           | 0.0693<br>(1.12) | −0.0128<br>(−1.11) | 0.0781<br>(9.77)    | 1                | 0.25<br>(0.618) | 0.0921<br>(2.25)   | −0.0210<br>(−2.82) | 0.0496<br>(15.16)   | 1                | 16.60<br>(0.000) | 1  |
| VR           | 0                | 0                  | −0.0302<br>(−10.23) | 3/2              | 7.08<br>(0.069) | 0                  | 0                  | −0.0153<br>(−12.09) | 3/2              | 37.22<br>(0.000) | 3  |
| CEV          | 0                | −0.0005<br>(−0.11) | −0.0821<br>(−2.45)  | 0.9944<br>(4.09) | 1.44<br>(0.230) | 0                  | −0.0043<br>(−1.36) | −0.1671<br>(−4.24)  | 0.3920<br>(2.99) | 5.97<br>(0.015)  | 1  |

The parameters are estimated by the classical GMM induced by the original orthogonality function  $h$  and by means of the RGMM induced by the modified orthogonality function  $h_t^{t,\tau}$  implied by a bound  $c=5.85$  on the influence function of the estimator for the unrestricted model;  $t$ -statistics are in parentheses. The value of Hansen's statistics ( $\xi$  for short) are reported with  $p$ -values in parentheses and associated degrees of freedom ( $df$ ). We use eurorates from February 1975 to April 2000 obtained from Datastream. The parameters are estimated from the following discrete-time system of equations:

$$r_t - r_{t-1} = \alpha + \beta r_{t-1} + \varepsilon_t$$

$$E(\varepsilon_t) = 0, \quad E(\varepsilon_t^2) = \sigma^2 r_{t-1}^{2\gamma}.$$

300 observations generated by a CIR model, the classical Hansen's test could not distinguish between a CIR and a Brennan Schwartz specification.

In order to further verify this inference we take, as usual, a closer look at the estimated robust weights that are given in Fig. 5.

We observe a cluster of influential observations around a similar period as for the US short rates. However, the Euro-DM patterns are slightly different, with the most influential points situated at the beginning and at the end of the October 1979–September 1982 sub-period, rather than in the middle in March 1980. This clustering of influential points corresponds to the regime found in Ang and Bekaert (2002b).

Additional influential observations are found at the beginning of 1975, coinciding with the end of the first OPEC crisis and following the adjustment period after the Bretton Woods era. The other three most influential points are observed at January 1987, July 1988 and December 1990. These observations almost coincide with three of the four largest specific events pointed out by Ball and Torous (1999) using a different approach: a Bundesbank announcement (December 1990), an OPEC production agreement (November 1988) and a tight credit market condition (December 1986). Moreover, when imposing

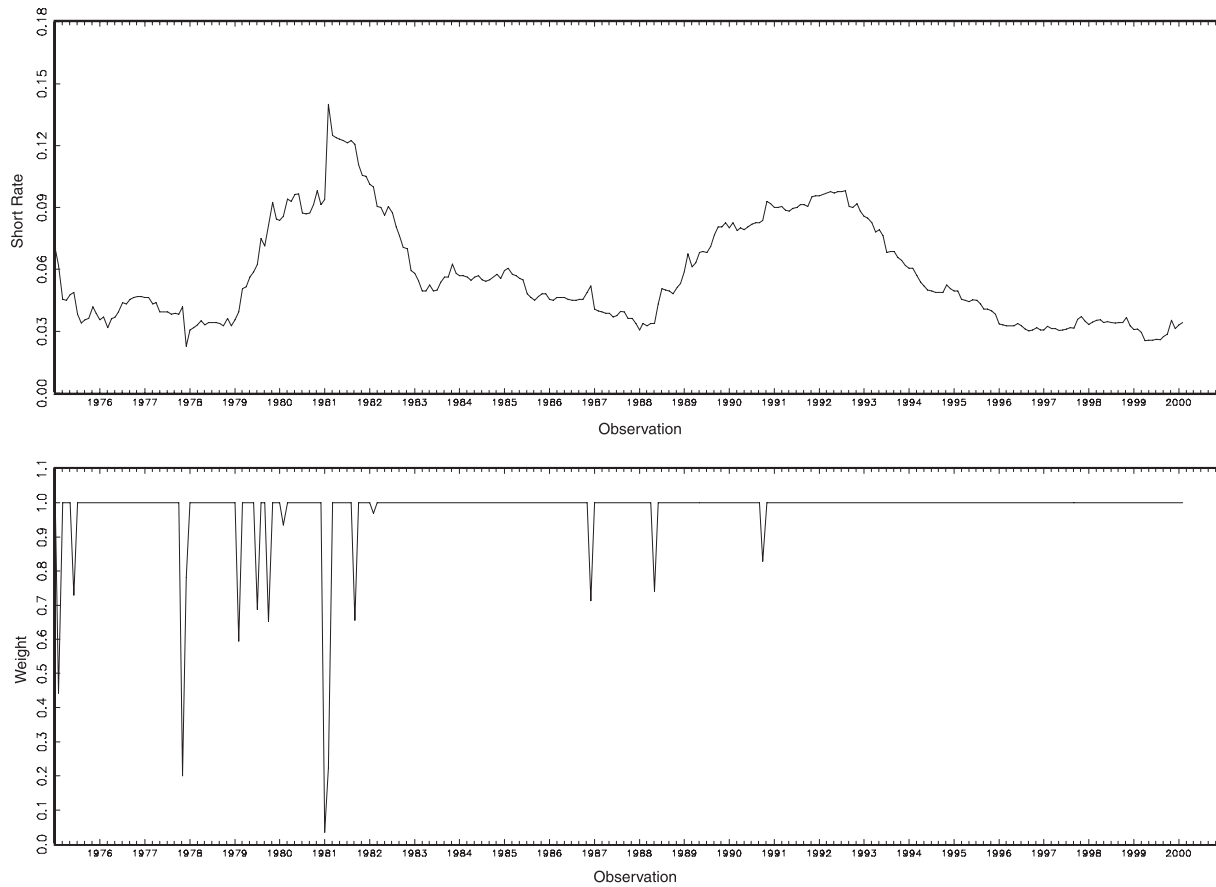


Fig. 5. Euro-mark data and weights implied by the RGMM estimator. Influential observations correspond to weights that are smaller than 1.

slightly more robustness ( $c$  between 5 and 5.5) we also observe an influential point at the ERM crisis in September 1992.

Finally, as for the US, in the period after 1982 only some isolated influential points are observed. This suggests again that a CIR model could be able to provide a satisfactory description of the short rate Euro-DM process starting in 1982.

#### **4. Conclusions**

In this paper, we have focused on the performance of the robust GMM for the selection of one-factor models for the short rate process within the standard CKLS framework and some of its extensions allowing for a nonlinear drift term and a structural break. We have found evidence of unstable GMM statistics causing unreliable classical GMM procedures and model selections. In contrast, RGMM model selection procedures perform well and offer a valid complement to the classical strategy. We have demonstrated how the new method uncovers general misspecification structures, thereby avoiding the danger of misleading statistical conclusions as in the classical GMM analysis. An important feature of the RGMM is the identification of general influential data points. In the CKLS dataseries, we find a cluster of influential points between 1979 and 1982, which is the well-known period of the Fed Experiment. This is not only true for the CKLS model, which is known to be misspecified when applied to US data spanning the 1979–1982 period, but also for a model with an additional quadratic drift term and for a model allowing for changing parameters during the 1979–1982 subperiod.

From the analysis above, we can learn about some of the risks of classical GMM testing procedures, where a model is tested against a more general nesting one. If the nesting model is possibly itself slightly misspecified, the resulting statistics may suggest misleading conclusions about the underlying data generating process. In such situations, the RGMM can help to highlight this problem using basically only some bounded, modified GMM orthogonality conditions of the original model. Before extending a model, it is therefore important to understand first whether the general structure of the data can support it. For instance, a robust methodology can show whether a more complex model is preferred just because of differences in a particular historical period. This is often relevant in particular in the context of pricing and hedging of financial instruments.

Taking into account the exceptional 1979–1982 period, an alternative strategy is to consider a simpler model like a CIR for the post 1982 period, as it is often done in practice (see [Zhang, 1999](#)). Similar results apply for Euro-DM data.

#### **Acknowledgements**

The authors would like to thank G. Bekaert (the editor), two referees, O. Scaillet and S. Cain for many helpful suggestions which improved the paper. Fabio Trojani gratefully acknowledges the financial support of the Swiss National Science Foundation (grant 1214-056679) and that of the NCCR FINRISK.

## References

- Ahn, D.H., Gao, B., 1999. A parametric nonlinear model of term structure dynamics. *Review Financial Studies* 12, 721–762.
- Aït-Sahalia, Y., 1996. Testing continuous-time models of the spot interest rate. *Review of Financial Studies* 9, 385–426.
- Ang, A., Bekaert, G., 2002a. Short rate nonlinearities and regime switches. *Journal of Economic Dynamics and Control* 26, 1243–1274.
- Ang, A., Bekaert, G., 2002. Regime switches in interest rates. *Journal of Business and Economics Statistics* 20, 163–182.
- Ball, C.A., Torous, W.N., 1995. Regime Shifts in Short Term Riskless Interest Rates, Working Paper UCLA 15–95.
- Ball, C.A., Torous, W.N., 1999. The stochastic volatility of short-term interest rates: some international evidence. *Journal of Finance* 54, 2339–2359.
- Bekaert, G., Hodrick, R.J., Marshall, D.A., 2001. Peso problem explanations for term structure anomalies. *Journal of Monetary Economics* 48, 241–270.
- Bliss, R.R., Smith, D.C., 1999. The elasticity of interest rate volatility: Chan, Karolyi, Longstaff, and Sanders revisited. *Journal of Risk* 1, 21–46.
- Brenner, J.B., Harjes, R.H., Kroner, K.F., 1996. Another look at models of the short-term interest rate. *Journal of Financial and Quantitative Analysis* 31, 85–107.
- Broze, L., Scaillet, O., Zakoian, J., 1995. Testing for continuous-time models of the short-term interest rate. *Journal of Empirical Finance* 2, 199–223.
- Cai, J., 1994. A Markov model of unconditional variance in ARCH. *Journal of Business and Economic Statistics* 12, 309–316.
- Chan, K.G., Karloyi, G.A., Longstaff, F.A., Sanders, A.B., 1992. An empirical comparison of alternative models of the short-term rate. *Journal of Finance* 47, 1209–1227.
- Dai, Q., Singleton, K.J., 2000. Specification analysis of affine term structure models. *Journal of Finance* 55, 1943–1978.
- Duffee, G., 1996. Idiosyncratic variation of treasury bill yields. *Journal of Finance* 51, 527–551.
- Gray, S., 1996. Modelling the conditional distribution of interest rates as a regime-switching model. *Journal of Financial Economics* 42, 27–62.
- Hampel, F.R., 1974. The influence curve and its role in robust estimation. *Journal of the American Statistical Association* 69, 383–393.
- Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J., Stahel, W.A., 1986. *Robust Statistics: The Approach Based on Influence Functions*. Wiley, New York.
- Hansen, L.P., 1982. Large sample properties of generalized method of moments estimators. *Econometrica* 50, 1029–1054.
- Heritier, S., Ronchetti, E., 1994. Robust bounded-influence tests in general parametric models. *Journal of the American Statistical Association* 89, 897–904.
- Huber, P., 1981. *Robust Statistics*. Wiley, New York.
- Jiang, G.J., 1998. Nonparametric modeling of U.S. interest rate term structure dynamics and implication on the prices of derivative securities. *Journal of Financial and Quantitative Analysis* 33, 465–497.
- Knez, P.J., Ready, M.J., 1997. On the robustness of size and book-to-market in cross-sectional regressions. *Journal of Finance* 52, 1355–1382.
- Koedijk, K.G., Nissen, F.G., Schotman, P.C., Wolff, C.P., 1997. The dynamics of short-term interest rate volatility reconsidered. *European Finance Review* 1, 105–130.
- McCulloch, J., Kwon, H., 1993. US Term Structure Data, 1947–1991. Working paper, Ohio State University.
- Nowman, K.B., 1997. Gaussian estimation of single-factor continuous time models of the term structure of interest rates. *Journal of Finance* 4, 1695–1706.
- Ronchetti, E., Trojani, F., 2001. Robust inference with GMM estimators. *Journal of Econometrics* 101, 37–69.
- Stanton, R., 1997. A nonparametric model of term structure dynamics and the market price of interest rate risk. *Journal of Finance* 52, 1973–2002.
- Zhang, H., 1999. Dynamics of the short-term interest rate after the 1979–1982 monetary experiment. *Journal of Fixed Income* 9, 35–41.